CLAIMS

What is claimed is:

1. A method for generating a statistic for phone lengths, with which the phone lengths can be controlled on the basis of this statistic during synthetic speech generation, comprising:

assigning phones of a spoken and recorded text that is segmented into phones, to phonemes of predetermined primary clusters composed of a plurality of phonemes, in each case one phone being assigned to a primary phoneme of one of the predetermined primary clusters if present in the spoken text in a context which is identical or similar to the context of the primary phoneme;

producing a primary statistic including at least an average phone length of all the phones assigned to a corresponding phoneme of one of the predetermined primary clusters;

assigning phones of the spoken and recorded text to phonemes of predetermined secondary clusters composed of phonemes, a number of phonemes of at least some secondary clusters differing from a number of phonemes of the predetermined primary clusters, in each case one phone being assigned to a secondary phoneme of one of the predetermined secondary clusters if present in the spoken text in a context which is identical to the context of the secondary phoneme; and

producing a secondary statistic including at least an average phone length of all the phones assigned to the secondary phoneme.

2. The method as recited in claim 1, wherein the number of phonemes of the primary clusters is constant.

3. The method as recited in claim 2, wherein the number of phonemes of the secondary clusters is variable, and the secondary clusters each include the phonemes of a word.

4. The method as recited in claim 3, wherein the primary statistic and the secondary statistic each includes a standard variation of a phone length.

5. The method for generating a statistic as claimed in claim 4, wherein the secondary statistic covers only selected secondary clusters whose frequency in the text is at least as large as a predetermined minimum frequency.

6. The method for generating a statistic as claimed in claim 5, wherein the minimum frequency is in the range from 3 to 10.

7. The method for generating a statistic as claimed in claim 6, wherein the phones are assigned to phonemes of the predetermined primary clusters using a predetermined list of phonemes grouped into the predetermined primary clusters, the phones being assigned to individual phonemes of the predetermined primary clusters in the list, and each individual association being stored.

8. The method as claimed in claim 7, wherein in each case the average phone length and the standard variation of the average phone length are calculated for the individual phonemes of the predetermined primary clusters in the list based on the individual associations that are stored.

9. The method as claimed in claim 1, wherein the phones are assigned to the phonemes of the predetermined secondary clusters using a predetermined list of phonemes grouped into the predetermined secondary clusters, the phones being assigned to individual phonemes of the predetermined secondary clusters in the list, and each individual association being stored.

10. The method as claimed in claim 9, wherein in each case the average phone length and the standard variation of the average phone length are calculated for the individual phonemes of the secondary clusters in the list on the basis of the stored associations.

11. . The method as recited in claim 2, wherein the number of phonemes in each of the predetermined primary clusters is equal to 3.

12. A method for determining a length of individual phones for speech synthesis, comprising:

calculating a primary statistic for phone lengths based on primary phonemes grouped into primary clusters and an average phone length assigned to the primary phonemes;

calculating a secondary statistic for phone lengths based on secondary phonemes grouped into secondary clusters and an average phone length assigned to the secondary phonemes;

determining whether a specified phoneme to be converted into speech and having a defined phone length has a corresponding phoneme in a respective secondary cluster;

assigning the average phone length of the secondary statistic to the corresponding phoneme in the respective secondary cluster if the specified phoneme matches the corresponding phoneme in the respective secondary cluster, and

assigning the average phone length of the primary statistic to a corresponding phoneme in a respective primary cluster if the specified phoneme does not match any phoneme in the secondary clusters.

13. A method for determining the length of the individual phones in speech synthesis as recited in claim 12 using the statistic generated by the method recited in claim 1.

14. A method as claimed in claim 12, wherein standard variations (G) of the average phone lengths (d') stored in the statistic are taken into account in determining the length (d) of the individual phones in accordance with the following formula

$$d = d' + G \cdot s,$$

where s is a speed scaling factor which is calculated according to the following formula

$$s = Rrel - 1,$$

Rrel being a ratio of speech speed to be spoken with respect to the speech speed with which the text on which the statistic is based has been spoken.

15. A device for generating a statistic for phone lengths to base control of the phone lengths during synthetic speech generation, comprising:

a computer system having a storage area in which a program for carrying out a method as recited in claim 1 is stored.

16. A device for determining the length of individual phones for speech synthesis, comprising:

a computer system having a storage area in which a program for carrying out a method as recited in claim 11 is stored.